Local structure in molecular complexes probed by multiple-scattering XAS

Andrea Di Cicco

INFM, LNF-INFN, Dipartimento di Fisica, Università degli Studi di Camerino, Via Madonna delle Carceri, 62032 Camerino (MC), Italy. E-mail: andrea.dicicco@unicam.it

The GNXAS (*n*-body distribution function X-ray absorption spectroscopy) method for multiple-scattering (MS) data analysis of EXAFS (extended X-ray absorption fine structure) data and the results recently obtained on molecular complexes relevant to biological matter are briefly reviewed and discussed. Practical MS calculations for important molecular fragments like Fe—O—O and Fe—C—N—Cu are presented in detail showing the potential of the techniques for measuring bond-angle distributions. The optimal conditions for obtaining accurate structural refinements using EXAFS measurements and modern data-analysis schemes are discussed as well as the current perspectives in the exploitation of the technique.

Keywords: molecular complexes; multiple-scattering X-ray absorption spectroscopy; biological materials; extended X-ray absorption fine structure.

1. Introduction

In recent times, the application of X-ray absorption spectroscopy (XAS) to the study of the local structure of molecular complexes relevant to biological matter has been shown to provide useful complementary information to crystallographic data. Successful XAS investigations using advanced multiple-scattering (MS) data-analysis methods have been performed by several groups on a variety of organic and inorganic complexes containing metal centers. In particular, the GNXAS method for XAS data analysis [*n*-body distribution function (GN) X-ray absorption spectroscopy (XAS); see Filipponi *et al.* (1995), Filipponi & Di Cicco (1995, 2000) and references therein] has been used to investigate the neighborhood of selected metal centers, fully exploiting the information about the local geometry contained in the multiple-scattering signal.

Some recent results include the application of such methods to the structural refinement of the Fe—Mo cofactor of nitrogenase (Nordlander *et al.*, 1993; Liu *et al.*, 1994; Conradson *et al.*, 1994), the determination of the N—O bonding in iron complexes (Westre *et al.*, 1994, 1995) and the study of the bond angle in cyanide-bridged Fe—Cu molecular assemblies relevant to cytochrome *c* oxidase (Zhang *et al.*, 1996, 1997). The GNXAS multiple-scattering method has also been used to study the local structure of organocobalt model compounds of vitamin B_{12} (Giorgetti *et al.*, 1996), the Cu enzyme galactose oxidase (Wang *et al.*, 1998), the active site of the tetanus neurotoxin (Meneghini & Morante, 1998) and hydroxocobalamin under different reduction conditions (Giorgetti *et al.*, 2000).

The availability of accurate crystallographic data can obviously limit the interest in using XAS for structure determination. However, there are some solid arguments that can stimulate the use of XAS as a suitable technique for biophysical studies. First, crystals of biological molecules suitable for crystallographic studies are not always available. Second, crystallographic refinements of very complex structures lead, not infrequently, to inaccurate atomic positions. Third, the local structure probed around selected atomic sites by XAS can provide information about bond and bond-angle distribution (including correlated vibrational terms) that is not included in crystallographic data. Moreover, it is not always clear whether the structure of a biological molecule is the same in crystalline form and in solution. XAS measurements can be carried out in both cases and the structure can be studied according to the aggregation state or chemical environment. In many interesting cases, accurate and unique information about the local structure of molecular fragments in biological matter can be obtained by XAS, provided that advanced data analysis is used on good quality experimental data.

2. Methodology

The reliability and efficiency of current methods for XAS data analysis has considerably improved in recent years due to major developments in theoretical approaches [see Gurman *et al.* (1984), Natoli & Benfatto (1986), Chou *et al.* (1987), Filipponi (1991), Tyson *et al.* (1992), Rehr & Albers (2000) and references therein]. In fact, the current standard in XAS data analysis is based on the comparison of accurate theoretical calculations with experimental data. Here I shall concentrate only on the description of one of the modern methods (GNXAS) used for analyzing the structural signal for high kinetic energies of the photoelectron (the so-called 'EXAFS' regime).

Current approximations for XAS are:

(i) The dipole approximation for the absorption matrix element, almost always valid in the typical available energy range for *K*- and *L*-edges associated with excitations of deep atomic levels.

(ii) The one-electron approximation, for which the many-body excitation process is described only by the transition of an optical electron from the core orbital into a continuum state. Many-body effects and in particular loss processes are accounted for by the imaginary part of the one-electron complex effective potential through the Hedin–Lundqvist approximation [see Chou *et al.* (1987), Tyson *et al.* (1992), Hedin & Lundqvist (1971) and references therein]. Other many-body effects associated with excitations of two or three electrons are presently introduced as empirical corrections [see Filipponi (1995) and references therein].

(iii) The muffin-tin approximation, where the photoelectron interacts with a collection of spherically symmetric non-overlapping potentials centered around each neighboring atom, embedded in a constant potential region. In common practice, the atomic potentials are obtained from spherical atomic self-consistent calculations and the bonding charge is accounted for by overlapping and averaging the neighbor atoms' charge density. It must be noted that more accurate non-muffin-tin atomic potentials are usually necessary for quantitative calculations of the near-edge structures (Foulis *et al.*, 1990, 1995; Joly *et al.*, 1999), while for higher photoelectron kinetic energies (starting from 20 eV above the edge) a muffin-tin approximation can be safely used.

The justification for the validity of the previous approximations stems from the weak XAS sensitivity to the low-energy details of the potential. The photoelectron is scattered by the deep spherical potential regions around the atomic nuclei. The core hole in the photoabsorber atom potential is important, and this is accounted for by means of an SCF (self-consistent field) atomic calculation for the atom in the excited electronic configuration. The XAS signal, defined as the modulation $\chi(k) = [\sigma(E) - \sigma_0(E)]/\sigma_0(E)$ of the X-ray absorption cross section $\sigma(E)$, originates from the scattering of the photoelectron wavefunction on the neighboring atomic potential centers. The modulation of the cross section is normalized to the atomic absorption cross section $\sigma_0(E)$ associated with the photo-

absorption process occurring without neighboring atoms. The $\chi(k)$ signal probes the local structure and is usually a regular oscillating function, where $k = [2m(E - E_e)]^{1/2}/\hbar$ is the modulus of the photoelectron wavevector (E_e being the threshold energy). MS effects are not negligible owing to the strong interaction of the secondary probe (the photoelectron) with the local charge density.

On the basis of these approximations, efficient data-analysis codes have been developed. In particular, the GNXAS method (Filipponi *et al.*, 1995; Filipponi & Di Cicco, 2000) possesses some unique features that make it particularly suitable for performing accurate XAS data analysis of ordered and disordered systems taking full account of MS effects. In the GNXAS approach, the interpretation of the experimental signal is based on a decomposition of $\chi(k)$ into a summation over irreducible n-body signals $\gamma^{(n)}(0, \ldots, i, j, k)$ (Filipponi & Di Cicco, 1995, 2000) associated with n-body atomic configurations, namely

$$\chi(0, i, j, \ldots, n) = \sum_i \gamma^{(2)}(0, i) + \sum_{(i,j)} \gamma^{(3)}(0, i, j) + \sum_{(i,j,k)} \gamma^{(4)}(0, i, j, k) + \cdots + \gamma^{(n)}(0, i, j, \ldots, n). \quad (1)$$

The $\gamma^{(n)}$ signals are associated with well defined atomic positions and can be calculated either directly, by means of continued fraction expansions (Filipponi, 1991), or by series summing of specific multiple scattering terms. The $\gamma^{(n)}(k)$ are usually regular oscillating functions of k:

$$\gamma^{(n)}(k) = A(k, \{r\}) \sin\big[kR_{\rm path} + \varphi(k, \{r\})\big]. \quad (2)$$

In (2), $kR_{\rm path}$ is the leading phase term, $R_{\rm path}$ being the length of the shortest path joining all of the atoms in the n-body configuration. The amplitude function $A(k, \{r\})$ depends on the order n, atomic types and geometry. Equation (1) refers to an ideal fixed atom configuration; in real experiments the XAS spectrum $\chi(k)$ is an averaged quantity that accounts for thermal and/or configurational disorder. Each sum in (1) is replaced by an integral over the appropriate n-body distribution function (g_n). For a single component system this is given by

$$\langle\chi(k)\rangle = \int_0^\infty {\rm d}r\ 4\pi r^2 \rho\ g_2(r)\ \gamma^{(2)}(r, k) + \int {\rm d}r_1 {\rm d}r_2 {\rm d}\theta\ 8\pi^2 r_1^2 r_2^2 \sin(\theta) \times \rho^2 g_3(r_1, r_2, \theta)\ \gamma^{(3)}(r_1, r_2, \theta, k) + \cdots. \quad (3)$$

The integrals formally extend to the whole coordinate space; however, because of the short-range nature of the $\gamma^{(n)}$ signals, due mainly to the exponential cutoff introduced by mean-free-path effects, they are actually limited to a region of linear dimensions of the order of a few Å. The dominant two-body term is usually accompanied by detectable three-body or even four-body high-frequency corrections that contain unique structural information about the short-range correlations in the system.

Within the GNXAS method, analysis of the structural signal is performed on the basis of (3) by refining a model structure (*i.e.* its local lowest-order distribution functions) and by comparing the corresponding $\chi(k)$ with experimental XAS. Proper account is taken of the configurational average, giving information about pair and higher-order correlations in ordered and disordered systems. The use of raw absorption data without any filtering makes it possible to perform a full statistical analysis of the results.

The GNXAS method (Filipponi & Di Cicco, 1995, 2000) has been successfully applied to several molecular, crystalline, nanophase, biological and disordered systems. Typical statistical errors are in the 0.001–0.01 Å range for the first-shell average bond distance depending on the system and on the quality of the experimental data.

Of course, especially for dilute biological specimens, the design of the experiment and the quality of the final XAS data are crucial aspects of the structural refinement. Good characterization of samples, low-temperature experiments and high statistics are required for accurate data analysis of complex systems using theoretical MS calculations. The following section reports some exemplary MS calculations of typical molecular fragments relevant to a wide class of inorganic and organic molecular complexes.

3. Examples of multiple-scattering calculations

3.1. Three-body terms: measuring bond-angle distributions

A diatomic molecule AB bonded to a metal center M represents a typical atomic configuration of interest for a variety of important organic and inorganic systems. For example, a study of the Fe—NO bond-angle distribution in different inorganic complexes is reported by Westre *et al.* (1994). Even in the presence of a complex molecular structure, the $M—A—B$ atomic configuration is associated with a well defined multiple-scattering signal that can be identified in the XAS spectra.

It is well known [see Westre *et al.* (1995) and references therein] that the amplitude of the MS signal increases for wider scattering angles and is particularly high for collinear configurations. The decomposition of the $M—A—B$ contributions into n-body terms results in a first-neighbor two-body $\gamma^{(2)}_{MA}$ signal, a second-neighbor $\gamma^{(2)}_{MB}$ signal and a single irreducible three-body term $\gamma^{(3)}$. The presence of the third atom B is therefore associated with an effective MS signal $\eta^{(3)}$ given by the two- and three-body contributions $\eta^{(3)} = \gamma^{(2)}_{MB} + \gamma^{(3)}$. The relevant coordinates R_{AM}, R_{AB} and θ_{BAM} for the three-atom molecular fragment are shown in Fig. 1. These are the natural coordinates associated with the chemical bonds and bond angles and define completely the three-body configuration.

The dominant contribution to the XAS spectrum is obviously the first-neighbor two-body $\gamma^{(2)}_{MA}$ signal, but the effective signal associated with the second neighbor (B) can be easily identified because of the different phase of the oscillating signal. In fact, as mentioned above, the leading term contributing to the phase of the $\gamma^{(n)}$ signals corresponds to the length $R_{\rm path}$ of the shortest scattering path ($R_{\rm path} \simeq 2R_{MA}$ for the first shell, $R_{\rm path} \simeq R_{MA} + R_{AB} + R_{BM}$ for the three-body term).

The interplay between the two-body and three-body contributions in the effective second-neighbor term can be better discussed by looking at a practical example. The MS irreducibile terms for a triangular Fe—O—O configuration (an oxygen molecule bonded to iron) have been calculated using Hedin–Lundqvist complex exchange-correlation potentials (Hedin & Lundqvist, 1971; Tyson *et al.*, 1992) and the muffin-tin approximation (GNXAS package).

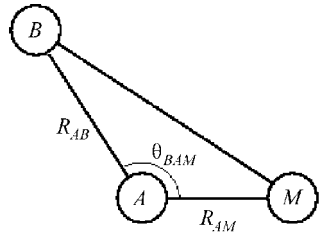

Figure 1
Coordinates of a triangular atomic configuration used for MS simulations.

Calculations were performed keeping the bond distances fixed ($R_{FeO} = 1.85$ Å and $R_{OO} = 1.15$ Å) over a broad range for the bond angle θ. Muffin-tin radii were chosen by suitable scaling of the atomic number radii ratios [Norman radii; see Norman (1976) and references therein] and were found to be $R_{MT}^{Fe} \simeq 1.14$ Å and $R_{MT}^{O} \simeq 0.66$ Å, giving an overlap of about 15% along the O—O bond [see also Westre *et al.* (1994, 1995) and references therein]. Calculations were performed using a 700 MHz laptop computer under Linux operating systems with 1 s of typical CPU time.

In Fig. 2 (upper panel), MS calculations for two different Fe—O—O triangular configurations (180° and 120°) are shown. The amplitude of the irreducible $\gamma^{(3)}$ signal is clearly enhanced in the collinear configuration, while the amplitude of the two-body signal is slightly higher at 120° because of the shorter second-neighbor distance [$\gamma_L^{(2)}$ curve, where L represents the longest bond of the triangle]. There is a clear change in the phase of the signals associated with the variation of the bond angle. The situation can be better understood by looking at the integrated amplitude of the signals as a function of the angle θ shown in Fig. 2 (lower panel). For each θ value, the amplitude of the MS irreducible signals has been averaged in the range 27.2–408 eV above the interstitial potential (taken as energy zero). The amplitude of the two-body $\gamma_L^{(2)}$ signal decreases smoothly, owing to the gradual elongation of the Fe—O second-neighbor distance. The amplitude of the three-body $\gamma^{(3)}$ signal shows the opposite behavior, with a clear maximum at about 180°. The effective second-neighbor $\eta^{(3)}$ is the sum of the aforementioned signals, and the angle dependence of its average amplitude (squares) depends critically on the phase difference between $\gamma^{(3)}$ and $\gamma_L^{(2)}$. In particular, the two signals are almost in antiphase around 120° and the average amplitude is flat around the minimum between 120° and 150°.

This result shows that the bond-angle can be easily determined for $\theta \geq 150°$. For any angular value, the $\gamma_L^{(2)}$ and $\gamma^{(3)}$ terms must be simultaneously considered as they are both important in determining the actual phase and amplitude of the effective $\eta^{(3)}$ signal. Clearly, in real vibrating systems the MS signals must be averaged using realistic distributions of distances and angles. The GNXAS fitting program (*fitheo*; see Filipponi & Di Cicco, 1995, 2000) performs such realistic averages using rigorous methods (Benfatto *et al.*, 1989) and Gaussian or non-Gaussian distributions. Triangular configurations with special symmetry (equilateral, isosceles, collinear) are properly accounted for within this approach.

The real situation of an organic or inorganic molecule is usually more complicated, as many other *n*-body terms must be considered besides the aforementioned MS $\eta^{(3)}$ signal. In optimal conditions (low-temperature and low-noise data), the correlation among different structural parameters can be minimized and quite accurate estimates of the angular distribution are feasible. As Westre *et al.* (1994) show, an accuracy of about 5° can be achieved even in the presence of a complicated molecular environment using modern data analysis and robust statistical evaluation of errors. In particular, the residual function is shown to increase up to one order of magnitude, changing the bond angle to the values corresponding to the minima [see Figs. 10 and 15 of Westre *et al.* (1994)].

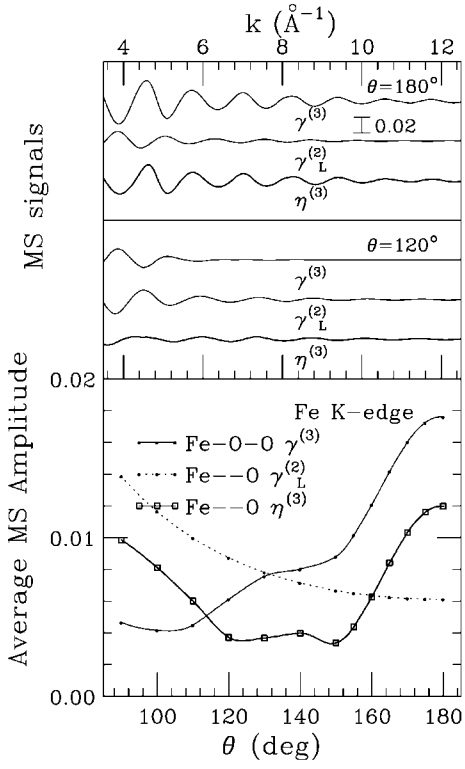

Figure 2
Upper panel: Fe *K*-edge irreducible $\gamma^{(2)}$ and $\gamma^{(3)}$ [$\eta^{(3)} = \gamma_L^{(2)} + \gamma^{(3)}$] MS signals calculated for $\theta = 180°$ (upper) and 120° (lower) Fe—O—O configurations. Lower panel: average amplitude of the $\gamma_L^{(2)}$ (dots, Fe—O), $\gamma^{(3)}$ (solid) and $\eta^{(3)}$ (squares) MS signals as a function of the bond angle θ.

3.2. Four-body terms: measuring correlated third-neighbor positions

A more complex structural situation, still quite frequent in a number of interesting systems, is when two metal sites are bridged by molecules like OH, O_2 or CN [see, for example, Zhang *et al.* (1996, 1997), Giorgetti *et al.* (1997)]. This molecular fragment can be regarded as a typical chain-like four-body configuration, like that shown in Fig. 3. In Fig. 3, the two metal sites *M*1 and *M*2 are bridged by a diatomic molecule *AB*. The four-body configuration is generally defined by six parameters, and a convenient choice is the use of the three shortest distances (chemical bonds R_{M1A}, R_{AB} and R_{BM2}), two bond angles (θ_{BAM1} and θ_{ABM2}) and the dihedral angle φ. A four-body configuration like the one depicted in Fig. 3 is usually associated with an important contribution to the XAS spectrum, especially when both bond angles approach 180° (super-collinear configuration).

The structural difference compared with the three-body configuration discussed above (see Fig. 1) is that there is an additional atom (namely *M*2 in Fig. 3) at a larger distance from atom *M*1 (thought of as the photoabsorbing center). Again, this four-body contribution can be easily identified because the phase of the oscillating signal associated with the presence of the atom *M*2 is different from those of the three-body *M*1—*A*—*B* and two-body *M*1—*A*, *M*1—*B* signals.

Several MS terms are associated with the presence of a fourth atom, like *M*2, in the structure. In particular, besides the irreducible four-body $\gamma^{(4)}$ term, there are two different three-body terms $\gamma_S^{(3)}$ (the first leg is 'short', *M*1—*A*—*M*2) and $\gamma_F^{(3)}$ (the first scattering atom is

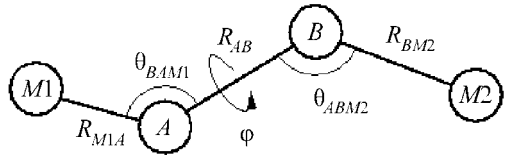

Figure 3
Coordinates of a four-body chain-like molecular fragment used for MS simulations.

'far', $M1-B-M2$). Of course, there is also the simple two-body $\gamma_L^{(2)}$ term ('long bond' $M1-M2$). When $M2$ is the photoabsorber and $M1$ is the fourth atom in the structure, the $\gamma_S^{(3)}$ signal refers to the $M2-B-M1$ configuration (the first is the shortest bond) and $\gamma_F^{(3)}$ to the $M2-A-M1$ configuration.

The effective third-neighbor signal $\eta^{(4)}$ for a chain-like configuration is defined as

$$\eta^{(4)} = \gamma_L^{(2)} + \gamma_S^{(3)} + \gamma_F^{(3)} + \gamma^{(4)}, \quad (4)$$

and in the following some practical calculations for the Fe—C—N—Cu structure shall be presented.

The MS irreducibile terms $\gamma_L^{(2)}$, $\gamma_S^{(3)}$, $\gamma_F^{(3)}$ and $\gamma^{(4)}$ have been calculated using the GNXAS package for both Fe and Cu K-edges using Hedin–Lundqvist complex exchange-correlation potentials (Hedin & Lundqvist, 1971; Tyson *et al.*, 1992) and the muffin-tin approximation (GNXAS program). Only the bond angles $\theta_1 = \theta_{FeCN}$ and $\theta_2 = \theta_{CNCu}$ were varied in the 90–180° range, while chemical-bond distances and the dihedral angle were kept fixed ($\varphi = 0°$, planar configuration, $R_{FeC} = 1.9$ Å, $R_{CN} = 1.1$ Å and $R_{CuN} = 1.9$ Å). Muffin-tin radii were $R_{MT}^{Fe} \sim 1.2$, $R_{MT}^{Cu} \sim 1.2$, $R_{MT}^{C} \sim 0.635$ and $R_{MT}^{N} \sim 0.635$ Å, giving an overlap of about 15% along the C—N bond [see also Zhang *et al.* (1996, 1997) and references therein]. Calculations were performed using a 700 MHz laptop computer under Linux operating systems with 30 s of typical CPU time for the four-body term.

Fig. 4 shows the Cu K-edge MS calculations for two different θ_2 angles (120° and 180°). The amplitude of the irreducible $\gamma^{(4)}$ and $\gamma^{(3)}$ signals is clearly enhanced in the super-collinear configuration $\theta_1 = \theta_2 = 180°$, while the amplitude of the two-body signal is slightly higher at 120° because of the shorter third-neighbor distance [$\gamma_L^{(2)}$ curve]. It is interesting to note that the four-body $\gamma^{(4)}$ signal is by far the most intense signal in the super-collinear configuration, but it is out of phase with both the $\gamma^{(3)}$ signals. The resulting four-body effective $\eta^{(4)}$ signal is therefore similar to the $\gamma^{(4)}$ signal but shows a much weaker amplitude. Conversely, for $\theta_2 = 120°$ the dominant contribution is the two-body $\gamma_L^{(2)}$ signal. The $\gamma_F^{(3)}$ term is slightly larger than $\gamma_S^{(3)}$ owing to the larger scattering angle found in the Cu—C—Fe configuration with respect to the Cu—N—Fe configuration.

A detailed analysis of the complex situation that can be found for various four-body structural configurations can be performed by looking at the integrated amplitude of the signals as a function of the angles θ_1 and θ_2, as shown in Fig. 5. In Fig. 5, for each angle, the amplitude of the MS irreducible signals has been averaged in the range 27.2–408 eV above the interstitial potential (taken as energy zero). The Cu and Fe K-edge XAS MS amplitudes are shown in the upper and lower panels, respectively. Calculations performed keeping $\theta_2 = 170°$ are shown in the left-hand panels, while those obtained keeping $\theta_1 = 180°$ are shown on the right. This set of calculations covers various structural four-body situations ranging from collinearity to large bond-angle distortions. Similarly to the three-body case (see §3.1), the amplitude of the two-body $\gamma_L^{(2)}$ term decreases smoothly with increasing bond angle owing to the elongation of the Fe—Cu third-neighbor distance. The amplitude of the three-body $\gamma^{(3)}$ signals displays a more complex behavior, showing clear maxima for configurations near to collinearity. The different dependences of the integrated amplitudes of the $\gamma_F^{(3)}$ and $\gamma_S^{(3)}$ signals on the selected bond angle is related to the actual scattering angle on the intermediate C or N atoms. As a rule of thumb, the larger the scattering angle the larger the integrated amplitude. For these three-body signals, the average contrast in amplitude is about 5 between collinear and far-from-collinear configurations. The behavior of the irreducible four-body $\gamma^{(4)}$ signal is simpler. Its amplitude shows a very marked maximum in the super-collinear configuration with a contrast of more than 10 with respect to the minimum. The angle dependence of the average amplitude (squares) of the effective third-neighbor $\eta^{(4)}$ term naturally

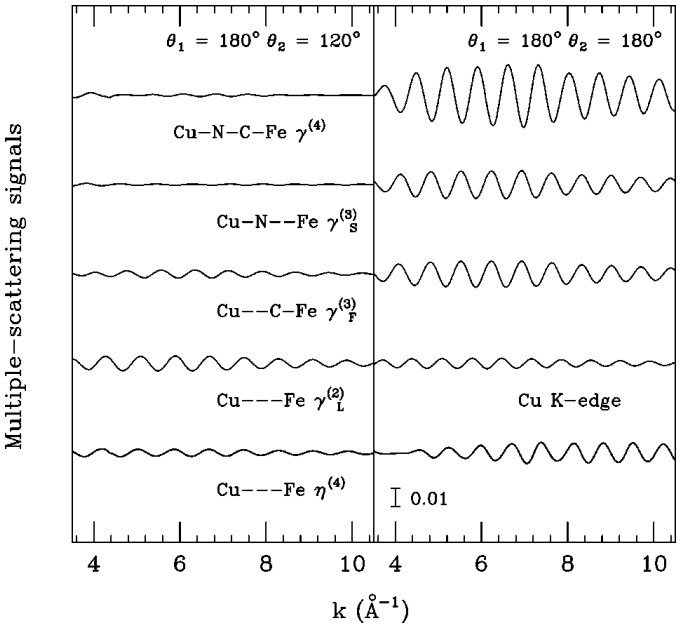

Figure 4
MS irreducible $\gamma^{(2)}$, $\gamma^{(3)}$ and $\gamma^{(4)}$ Cu K-edge XAS signals associated with a four-body chain-like Fe—C—N—Cu configuration. The total effective four-body signal $\eta^{(4)}$ is given by the sum of the n-body signals involving the fourth atom of the fragment (Fe). The results of the calculations for $\theta_2 = \theta_{CNCu} = 120°$ and 180° are reported in the left-hand and right-hand panels, respectively.

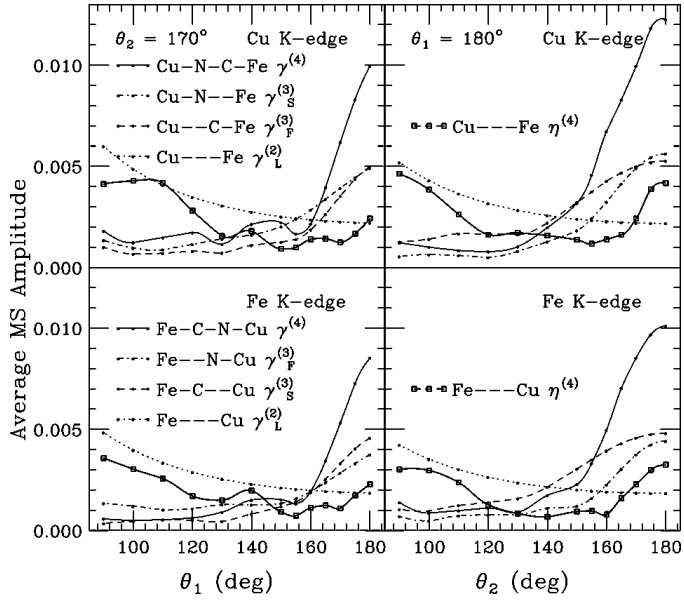

Figure 5
The four panels contain the average amplitudes of the MS signals associated with the four-body Fe—C—N—Cu configuration as a function of the bond angles. Fe and Cu K-edge calculations are reported in the lower and upper panels, respectively. The left-hand panels contain the average amplitudes obtained for fixed $\theta_2 = 170°$ while the right-hand panels refer to calculations with $\theta_1 = 180°$. All of the two-body $\gamma_L^{(2)}$ (dots), three-body $\gamma_F^{(3)}$ and $\gamma_S^{(3)}$ (dashed and dot-dashed), four-body $\gamma^{(4)}$, and total $\eta^{(4)}$ (squares) MS signals are reported.

depends critically on the phase differences between the $\gamma^{(4)}$, $\gamma^{(3)}$ and $\gamma_L^{(2)}$ signals. The amplitude of the $\eta^{(4)}$ contribution is quite weak in the 120–160° bond-angle range, with typical amplitudes of 0.001–0.002 (detectable in XAS spectra of 10^4 signal-to-noise ratios). This is valid when the second bond angle, either FeCN or CuNC, is almost collinear. Owing mainly to the irreducible four-body $\gamma^{(4)}$ contribution, a marked enhancement of the effective $\eta^{(4)}$ signal is obtained for quasi-collinear contributions with a maximum contrast of about 4.

These features were used to measure the bond-angle distribution in the cyanide-bridged $[(py)(OEP)Fe–CN–Cu(Me\_6tren)]^{2+}$ and $[(py)(OEP)Fe–CN–Cu(TIM)]^{2+}$ complexes (Zhang *et al.*, 1997) and to detect four-body Fe—CN—Co signals in mixed hexacyanoferrates (Giorgetti *et al.*, 1997). Generally speaking, bond-angle distributions associated with chain-like four-body configurations can be determined using good low-noise XAS data for average bond angles $\theta \geq 160°$. The whole third-neighbor contribution $\eta^{(4)}$ must be properly calculated and accounted for by averaging the signal with a realistic distribution of distances and angles, as implemented in the GNXAS fitting program (*fitheo*; see Filipponi & Di Cicco, 1995, 2000).

4. Conclusion and perspectives

The use of XAS for investigating the local structure of molecular complexes relevant to biological matter has been shown to give useful and reliable information about distance and angle distributions around selected atomic species. Previous applications of modern data-analysis methods based on accurate multiple-scattering simulations have already shown that structural results obtained by XAS nicely complement those available using other techniques [see, for example, Nordlander *et al.* (1993), Conradson *et al.* (1994), Westre *et al.* (1994, 1995), Zhang *et al.* (1996, 1997), Giorgetti *et al.* (1996, 2000), Wang *et al.* (1998), Meneghini & Morante (1998), and references therein].

In particular, the GNXAS method for data analysis, briefly described in this work, has been shown to provide accurate information about bond distances and angles in several systems containing well defined molecular fragments around metal centers. The decomposition into *n*-body terms typical of that method is shown to be particularly useful for obtaining robust information about the local two-body, three-body and even four-body distributions. The unique sensitivity to the short-range structure and the presence of important multiple-scattering contributions are shown to be key factors for the interpretation of the results of the XAS technique.

In this paper, specific examples of MS calculations for exemplary three- and four-body configurations containing one or more metal centers are explicitly discussed. The sensitivity of the technique to the presence of distant neighbors is discussed using the irreducible *n*-body MS γ^n terms. The complex interplay between different *n*-body terms related to the same structure is discussed in detail, showing that direct measurements of the bond-angle distributions are feasible for typical configurations with average bond angles $\theta \geq 150$–160°. Direct calculations of the three- and four-body terms in a broad range of bond angles for important molecular fragments such as Fe—O—O and Fe—C—N—Cu are presented for the first time.

The results of the XAS investigations presented and mentioned in this paper show that modern data analysis can be used to give unique information about the local structure in molecular complexes. Information about bond-angle distribution can be obtained from low-noise (signal-to-noise ratio 10^4) XAS data possibly obtained at low temperatures in order to increase the amplitude of the structural signal. The present scheme for data analysis can be safely used in a number of interesting cases, but the presence of the complex molecular environments that frequently occur in real systems clearly complicates the determination of specific structural features. Thus, an important advance in using MS calculations would be the classification of molecular fragments occurring more frequently in biological matter, like imidazole and porphyrin rings, in order to devise suitable parameterization schemes to overcome the simple application of the *n*-body expansion, originally thought to be used for simpler structures. MS XAS contributions related to those complex substructures can be calculated as a whole, classified in an appropriate way and successively used for structural refinements in real biological systems. This clearly will require efforts both at the theoretical and methodological level but could greatly simplify the data analysis in real systems and improve the quality of the refinements. Further advances in the interpretation of XAS data of complex systems could be achieved along these lines.

I am indebted to K. O. Hodgson, B. Hedman and C. R. Natoli for introducing me to this interesting subject and for their support in this activity.

References

Benfatto, M., Natoli, C. R. & Filipponi, A. (1989). *Phys. Rev. B*, **40**, 9626.
Chou, S., Rehr, J. J., Stern, E. A. & Davidson, E. R. (1987). *Phys. Rev. B*, **35**, 2604.
Conradson, S. D., Burgess, B. K., Newton, W. E., Di Cicco, A., Filipponi, A., Wu, Z. Y., Natoli, C. R., Hedman, B. & Hodgson, K. O. (1994). *Proc. Natl Acad. Sci.*, **91**, 1290.
Filipponi, A. (1991). *J. Phys. Condens. Matter*, **3**, 6489.
Filipponi, A. (1995). *Physica B*, **208/209**, 29.
Filipponi, A. & Di Cicco, A. (1995). *Phys. Rev. B*, **52**, 15135.
Filipponi, A. & Di Cicco, A. (2000). *Task Q.* **4**, 575.
Filipponi, A., Di Cicco, A. & Natoli, C. R. (1995). *Phys. Rev. B*, **52**, 15122.
Foulis, D. L., Pettifer, R. F., Natoli, C. R. & Benfatto, M. (1990). *Phys. Rev. A*, **41**, 4682.
Foulis, D. L., Pettifer, R. F. & Sherwood, P. (1995). *Europhys. Lett.* **29**, 647.
Giorgetti, M., Ascone, I., Berrettoni, M., Conti, P., Zamponi, S. & Marassi, R. (2000). *J. Biol. Inorg. Chem.* **75**, 156.
Giorgetti, M., Berrettoni, M., Conti, P., Di Cicco, A., Marassi, R. & Ascone, I. (1996). *Organometallics*, **15**, 3491.
Giorgetti, M., Berrettoni, M., Filipponi, A., Kulesza, P. J. & Marassi, R. (1997). *Chem. Phys. Lett.* **275**, 108.
Gurman, S. J., Binsted, N. & Ross, I. (1984). *J. Phys. C*, **17**, 143.
Hedin, L. & Lundqvist, B. I. (1971). *J. Phys. C*, **4**, 2064.
Joly, Y., Cabaret, D., Renevier, H. & Natoli, C. R. (1999). *Phys. Rev. Lett.* **82**, 2398.
Liu, H. I., Filipponi, A., Gavini, N., Burgess, B. K., Hedman, B., Di Cicco, A., Natoli, C. R. & Hodgson, K. O. (1994). *J. Am. Chem. Soc.* **116**, 2418.
Meneghini, C. & Morante, L. (1998). *Biophys. J.* **75**, 1953.
Natoli, C. R. & Benfatto, M. (1986). *J. Phys.* (*Paris*) *Colloq. C*8, **47**, 11.
Nordlander, E., Lee, S. C., Cen, W., Wu, Z. Y., Natoli, C. R., Di Cicco, A., Filipponi, A., Hedman, B., Hodgson, K. O. & Holm, R. H. (1993). *J. Am. Chem. Soc.* **115**, 5549.
Norman, J. G. (1976). *Mol. Phys.* **31**, 1991.
Rehr, J. J. & Albers, R. C. (2000). *Rev. Mod. Phys.* **72**, 621.
Tyson, T. A., Hodgson, K. O., Natoli, C. R. & Benfatto, M. (1992). *Phys. Rev. B*, **46**, 5997.
Wang, Y., Bois, J. L. D., Hedman, B., Hodgson, K. O. & Stack, T. D. P. (1998). *Science*, **279**, 537.
Westre, T. E., Di Cicco, A., Filipponi, A., Natoli, C. R., Hedman, B., Solomon, E. I. & Hodgson, K. O. (1994). *J. Am. Chem. Soc.* **116**, 6757.
Westre, T. E., Di Cicco, A., Filipponi, A., Natoli, C. R., Hedman, B., Solomon, E. I. & Hodgson, K. O. (1995). *J. Am. Chem. Soc.* **117**, 1566.
Zhang, H. H., Filipponi, A., Di Cicco, A., Lee, S. C., Scott, M. J., Holm, R. H., Hedman, B. & Hodgson, K. O. (1996). *Inorg. Chem.* **35**, 4819.
Zhang, H. H., Filipponi, A., Di Cicco, A., Scott, M. J., Holm, R. H., Hedman, B. & Hodgson, K. O. (1997). *J. Am. Chem. Soc.* **119**, 2470.